Psychology 4111: Assignment 1

Worth: 5% of final grade
Due Date: Midnight, 25-Oct-2021 (electronic submission via D2L)

Part A (21 points)

Instructions. In the lecture slides on hypothesis testing, the first problem involved flipping a coin N=11 times, and using the decision rule that minimizes the overall probability of error. In that example, H_0 stated that p, the probability of a Head on each toss, was 0.5; and H_1 stated that $p = 0.15$. For the current example, $N = 25$, H_0 states that $p = 0.5$, and H_1 states that $p = 0.75$. **Your task is to modify the syntax** in file *A1-PartA-2021-TEMPLATE.sps* to make it carry out the computations needed for the current problem. The parts of the syntax that need to be replaced are enclosed in square brackets. For example, you need to write the code when it says: **[INSERT YOUR COMMANDS HERE]**.

The breakdown of marks assigned to each command is as follows:

1. * Issue NEW FILE and DATASET CLOSE commands to start with a new empty dataset.	[2]
2. * Create variable X = the current row number minus 1.	[1]
3. * Format X to show no decimals.	[1]
4. * Compute new variables pX0 and pX1 where pX0 = p(X | H0) and pX1 = p(X | H1).	[2]
5. * Format pX0 and pX1 to display 4 decimals then list the data.	[3]
6. * Plot the two binomial distributions as bar charts.	[2]
7. * Compute a new variable called Reject, and set it equal 1 on rows where you would reject H0 and 0 on rows where you would not reject H0.	[1]
8.* Format variable reject to show only 1 decimal.	[1]
9. *Add value labels "Reject H0" and "Fail to reject H0" to variable Reject.	[1]
10. * List the data.	[1]
11. * Split the file by variable Reject, then use the DESCRIPTIVES command to display the SUMS of variables pX0 and pX1.	[3]
12. * Fill in alpha, beta, and power in the comment lines on the syntax window.	[3]

Part B (29 points)

Instructions. In the lecture slides on hypothesis testing, the first example using a non-directional (or two-tailed) H_1 involved flipping a coin N=13 times, and testing the null hypothesis that the coin was fair against the two-tailed alternative that it is unfair. In this part of the assignment, you will carry out computations for a similar example, but with N = 30 tosses of the coin. **Your task is to modify the syntax** in file *A1-PartB-2021-TEMPLATE.sps* to make it carry out the needed computations. As in Part A, the parts of the syntax that need to be replaced are enclosed in square brackets.

The breakdown of marks assigned to each command is as follows:

1. * Issue two commands needed to start with a new empty dataset.	[2]
2. * Create variable X = the current row number minus 1.	[1]
3. *Format X to show no decimals.	[1]
4. * Compute X = the p(X|H0) for the value of X on each row.	[1]
5. * Compute LeftTail = the cumulative probability of X in the lower (left) tail of the distribution.	[1]
6. * Compute RightTail = the right (upper) tail cumulative probability of X.	[2]
7. * Format variables pX to RightTail to show 8 decimals. List.	[2]
8.* Add appropriate variable labels to variables pX, LeftTail and RightTail.	[3]
9. * Plot the binomial distribution as a bar chart.	[1]
10. *Plot both the left and right tail cumulative probabilities in one (multiple) line graph.	[3]
11. * Create variable Reject05 = 1 on rows where you would reject H0 and 0 on rows where you would fail to reject H0.	[1]
12. * Create variable Reject01 = 1 on rows where you would reject H0 and 0 on rows where you would fail to reject H0.	[1]
13. * Format variables Reject05 and Reject01 to show no decimals, then list the data in the output window.	[2]
14. * Use a TEMPORARY SELECT IF followed by a DESCRIPTIVES command to display the SUM of the pX values for records where Reject05 = 1.	[2]
15. * Fill in the decision rule and the actual alpha value on the following comment lines.	[2]
16. * Use a TEMPORARY SELECT IF followed by a DESCRIPTIVES command to display the SUM of the pX values for records where Reject01 = 1.	[2]
17. * Fill in the decision rule and the actual alpha value on the following comment lines.	[2]

Part C (38 Points)

Instructions. Wagner & colleagues (1988) investigated the possibility that incoming college students who experience more stressful events will exhibit more symptoms. Data file *Wagner_hassles.csv* contains variable HASSLES, the number of stressful events experienced in the last month. (The *.csv* extension stands for *comma-separated-variable*. This means that the data fields within the file are separated by commas. If you double-click on a .csv file, it will open in Excel.) Data file *Wagner_symptoms.txt* is a text file containing variable SYMPTOMS, the number of symptoms (from the Hopkins Symptom Checklist) experienced in the last month. (If you double-click on a *.txt* file, it will open in the default text editor, usually Notepad. The symptoms data file also has a field delimiter, but it is not a comma.) Both data files contain an ID variable.

Write (and save) SPSS syntax to carry out the following:

1. Issue NEW FILE and DATASET CLOSE ALL commands to start with a clean slate.	[2]
2. Use a FILE HANDLE command to define MyFolder as the folder in which you have stored the data files.	[1]
3. Using the file handle you defined in step 2, open *Wagner_hassles.csv*, and assign it the dataset name *hassles*. (**Hint**: A *.csv* file is a special kind of text file.)	[2]
4. Using the file handle you defined in step 2, open *Wagner_symptoms.csv*, and assign it the dataset name *symptoms*	[2]
5. Merge the two datasets, and call the resulting dataset *Wagner*. Then close the *hassles* and *symptoms* datasets.	[3]
6. Add variable labels for **hassles** and **symptoms**.	[1]
7. Create histograms for variables **hassles** and **symptoms**.	[1]
8. Create a scatter-plot via the *legacy dialogs* with X = **hassles** and Y = **symptoms**. Use the Chart Editor to add a regression line (or fit line) to the graph.	[3]
9. Use the Chart Builder dialog to generate a scatter-plot with X = hassles and Y = symptoms. See the ***Simple Scatterplot with Fit Line*** example found under GGRAPH in the Command Syntax Reference manual. (You'll find the syntax reference manual in the SPSS Help menu.) Use that example to modify your GGRAPH syntax so that a regression line is added to the plot.	[5]
10. Estimate a simple linear regression model with X = **hassles**, Y = **symptoms**. Ask for confidence intervals on the estimates. Save the unstandardized predicted values & residuals .	[5]
11. Produce a residual plot with X = fitted value of hassles and Y = residual.	[2]
12. Compute the correlation between fitted value of hassles and the residual.	[1]
13. Write a brief summary of the results (as you would do in your thesis), and comment on what the results mean.	[10]

PLEASE NOTE THAT QUESTION 13 MUST BE DONE INDEPENDENTLY!